R E P O R T  R E S U M E S

ED 014 584 VT 003 679

TECHNICAL REPORT ON COMPARABILITY AND RELIABILITY OF SCORES ON USES AND U.S. CIVIL SERVICE COMMISSION TYPING TESTS.

BUREAU OF EMPLOYMENT SECURITY, WASHINGTON, D.C.

PUB DATE AUG 65

EDRS PRICE MF-$0.25 HC-$0.44 9P.

DESCRIPTORS- *TYPEWRITING, *TEST RELIABILITY, *COMPARATIVE TESTING, *TESTS,

THIS STUDY WAS CONDUCTED TO (1) DETERMINE WHETHER DIFFERENCES IN TYPING TEST LENGTH, CONTENT, FORMAT, AND TIME LIMIT HAVE ANY EFFECT ON TEST SCORES, AND (2) COMPARE THE RELIABILITY OF 5-MINUTE TYPING TESTS WITH THAT OF 10-MINUTE TESTS. TWO EQUIVALENT FORMS OF THE U.S. EMPLOYMENT SERVICE TYPING TEST AND ONE FORM OF THE U.S. CIVIL SERVICE COMMISSION TYPING TEST WERE USED IN THE STUDY. THE SAMPLE CONSISTED OF 235 APPLICANTS FOR TYPING AND STENOGRAPHIC POSITIONS TESTED IN FOUR LOCAL OFFICES OF THE OHIO STATE EMPLOYMENT SERVICE AND 248 APPLICANTS TESTED IN FOUR LOCAL OFFICES OF THE WASHINGTON STATE EMPLOYMENT SERVICE. SCORES ON 5-MINUTE ADMINISTRATIONS WERE FOUND TO BE COMPARABLE TO THOSE ON 10-MINUTE ADMINISTRATIONS. THE RELIABILITY OF SPEED SCORES WAS FOUND TO BE THE SAME FOR 10-MINUTE TESTS AS FOR 5-MINUTE TESTS, BUT THE RELIABILITY OF ERROR SCORES WAS SOMEWHAT HIGHER FOR 10-MINUTE TESTS THAN FOR 5-MINUTE TESTS. NO SIGNIFICANT DIFFERENCES WERE FOUND BETWEEN THESE TESTS IN EITHER SPEED SCORES OR ERROR SCORES. (PS)

ED014584

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

# TECHNICAL REPORT

# ON

# COMPARABILITY AND RELIABILITY OF SCORES ON USES AND U.S. CIVIL SERVICE COMMISSION TYPING TESTS

VT003679

U. S. Employment Service
in Cooperation with
Ohio and Washington State Employment Services

August 1965

# Comparability and Reliability of Scores on USES and U. S. Civil Service Commission Typing Tests


## Summary

This study was conducted to (a) determine whether differences in typing test length, content, format and time limit have any effect on test scores, and (b) compare the reliability of 5-minute typing tests with that of 10-minute tests. Two forms of the USES typing test and one form of the U. S. Civil Service Commission typing test were used in this study. No significant differences were found between these tests in either speed scores or error scores. Scores on 5-minute administrations were found to be comparable to those on 10-minute administrations. The equivalent form reliability of speed scores was found to be the same for 10-minute tests as for 5-minute tests, but the reliability of error scores was somewhat higher for 10-minute tests than for 5-minute tests.

## I. Background

Typing tests are used extensively by a variety of organizations, including many of the employers to which the USES makes referrals for typing positions. Except within large organizations such as the USES, typing tests are usually not standardized, but vary as to length, type of material used in the text, format, and time limit. The primary purpose of this study was to determine whether these variations in typing tests are associated with differences in the scores obtained on the tests. If it was determined that the kind of test taken had no significant effect on an individual's score, employers could accept typist applicants tested by the USES without further testing.

This study also provided an opportunity to compare the reliability of 5-minute typing tests with that of 10-minute tests. A previous study[1] had shown that the current 5-minute USES typing test forms have a median alternate form reliability of .96 for speed and .72 for errors. The effect on reliability of a longer administration time was not determined in that study.

## II. Objectives

A. To determine whether differences in typing test length, content, format and time limit have a significant effect on speed or error scores.

B. To compare the reliability of speed and error scores on a 5-minute typing test with scores on a 10-minute test.

## III. Method

### A. Sample

The sample consisted of 235 applicants for typing and stenographic positions tested in four local offices of the Ohio State Employment Service, and 248 applicants tested in four local offices of the Washington State Employment Service. The individuals from each office were alternately assigned to two groups, I and II, at the time of testing. The characteristics of age, education, and work experience of the individuals in each group are summarized in Table 1. Group I contained 22 males and 220 females. Group II contained 21 males and 220 females.

---

[1] U. S. Department of Labor, Bureau of Employment Security. Technical Report on Modernization of USES Typing Tests. Washington, D. C.: November 1964.

B. Tests

Forms A and B of the USES typing test were used in this study. Like all six of the USES typing test forms, these two forms have been shown to be equivalent in difficulty and have the same characteristics: i.e., their length is about 400 words, the text is typical business correspondence in the first or second person, the format is three one-paragraph letters, and the normal time limit is 5 minutes.

The typing test used for comparison was a form of the U. S. Civil Service Commission (CSC) typing test. It differs from the USES forms in several respects: The text is only about half as long; the content is of a general, expository and impersonal nature; it is in ordinary paragraph format; the time limit is 10 minutes. (While this study was in progress, the U. S. Civil Service Commission released new typing tests with a 5-minute time limit).

One point of similarity between the USES forms and the CSC form is that the average number of strokes per word is approximately the same: 6.0 for the USES test, and 6.1 for the CSC test.

C. Test Administration and Scoring

The plan of test administration was as follows:

| Group I | Group II |
|---|---|
| 1st test-CSC, 10 minutes | 1st test-USES Form B, 10 minutes |
| Rest period, 5 minutes | Rest period, 5 minutes |
| 2nd test-USES Form A, 10 minutes | 2nd test-USES Form A, 10 minutes |

At the end of 5-minutes of typing on each test, the examinees were instructed to mark their test papers to indicate their progress at that point, then to immediately continue typing.

Scoring was carried out in accordance with the directions in the USES Guide to the Use of Typing, Dictation and Spelling Tests. These directions are also very similar to those used by the CSC. A speed score (gross words per minute) and an error score were obtained for the first 5 minutes of each test and for the full 10 minutes of each test. Error scores on the 5-minute segments were doubled so they would be directly comparable to the error scores on the full 10-minute tests. A net words per minute score was also computed for each test at each length. Net scores were obtained by subtracting the number of errors from the gross words per minute score.

## IV. Results

### A. Effects of Variations in Test Content on Test Scores

Table 2 shows the mean and standard deviation of each of the six scores on the USES and CSC typing tests taken by Group I.

The differences between mean scores on the two tests are too small to be of practical importance. The difference between the two tests on mean gross words per minute was only 0.4 words-per-minute at the 5- and 10-minute lengths. The difference in mean errors between the two tests was less than one error at the 5-minute length, and only 0.3 error at the 10-minute length. Significance tests were made of differences between the means for each of the six pairs of corresponding scores. None of these differences reached significance at the .05 level. The differences between standard deviations of scores on the two tests were also small, and none was statistically significant.

### B. Effects of Differences in Test Time Limits on Test Scores

Table 3 shows the means and standard deviations of each of the six scores on the two USES typing test forms taken by Group II.

Significance tests were made of the differences between means of the five-minute and ten-minute administration scores for each of the six pairs of means. None of these differences was statistically significant.

### C. Effects of Differences in Test Content and Time Limit on Test Reliability

Table 4 shows data on reliability of scores. It can be seen that the correlations between the USES and CSC tests are almost identical to the correlations between two forms of the USES test at each time limit and for each of the three scores. Thus, variations in content of the USES and CSC tests do not affect the correlation between scores on the two tests.

The test-retest reliability of gross words per minute scores for 5-minute administrations is about the same as the reliability for 10-minute administrations. There is a difference, however, between the reliability of 5-minute error scores and 10-minute error scores. There is also some difference between the reliability of 5- and 10-minute net words per minute scores.

Since the reliabilities of error scores between 5-minute administrations found in this study were lower than those obtained in some previous

TABLE I

MEANS AND STANDARD DEVIATIONS OF AGE, EDUCATION, AND WORK EXPERIENCE

| Group | Age (years) | | Education (years) | | Experience (Months) | |
|---|---|---|---|---|---|---|
| | Mean | S.D. | Mean | S.D. | Mean | S.D. |
| I (N = 242) | 29.2 | 11.4 | 12.5 | 1.3 | 42.1 | 66.0 |
| II (N = 241) | 27.0 | 10.8 | 12.4 | 1.0 | 41.0 | 71.8 |

TABLE 2

MEANS AND STANDARD DEVIATIONS OF SCORES ON CSC AND USES TYPING TEST FORMS AT FIVE- AND TEN-MINUTE LENGTHS

(GROUP I: N = 242)

| | Gross Words per minute | | Errors* | | Net Words per minute | |
|---|---|---|---|---|---|---|
| | Mean | S.D. | Mean | S.D. | Mean | S.D. |
| First 5 Minutes | | | | | | |
| CSC Form | 47.0 | 10.7 | 17.7 | 11.8 | 29.3 | 17.4 |
| USES Form A | 47.4 | 11.0 | 16.8 | 10.8 | 30.6 | 17.1 |
| Full 10 Minutes | | | | | | |
| CSC Form | 46.3 | 10.8 | 17.8 | 11.1 | 28.5 | 17.4 |
| USES Form A | 46.7 | 10.9 | 17.5 | 10.6 | 29.2 | 16.6 |

*Error scores shown for 5 minutes are double the actual obtained scores

TABLE 3

MEANS AND STANDARD DEVIATIONS OF SCORES
ON USES TYPING TEST FORMS AT
FIVE- AND TEN-MINUTE LENGTHS

(GROUP II: N = 241)

| | Gross Words per minute | | Errors* | | Net Words per minute | |
|---|---|---|---|---|---|---|
| | Mean | S.D. | Mean | S.D. | Mean | S.D. |
| USES Form B | | | | | | |
| First 5 Minutes | 46.1 | 11.4 | 17.2 | 11.5 | 28.8 | 17.3 |
| Full 10 Minutes | 45.2 | 11.2 | 18.2 | 11.2 | 27.0 | 17.2 |
| USES Form A | | | | | | |
| First 5 Minutes | 47.1 | 11.3 | 17.9 | 11.4 | 29.3 | 17.1 |
| Full 10 Minutes | 46.4 | 11.0 | 19.0 | 11.1 | 27.4 | 16.9 |

*Error scores shown for 5 minutes are double the actual obtained scores.

TABLE 4

PRODUCT-MOMENT CORRELATIONS BETWEEN SCORES ON USES AND
CSC TYPING TESTS AND SCORES ON TWO USES TYPING
TEST FORMS AT FIVE- AND TEN-MINUTE LENGTHS

| | Gross Words per Minute | Errors | Net Words per Minute |
|---|---|---|---|
| 5-Minute Length | | | |
| CSC - USES Form A (Group I) | .95 | .66 | .84 |
| USES Form B - USES Form A (Group II) | .96 | .66 | .84 |
| 10-Minute Length | | | |
| CSC - USES Form A (Group I) | .98 | .80 | .91 |
| USES Form B - USES Form A (Group II) | .97 | .79 | .90 |

unpublished USES studies, an additional analysis was made of the data. In the previous studies the reliability coefficients were computed for two 5-minute tests administered in succession, whereas the coefficients in Table 4 are for two 5-minute tests separated in time by an additional 5-minute test segment and a 5-minute rest period. In order to obtain coefficients on successive 5-minute tests for the present data, error scores were computed separately for the two 5-minute halves of the first test administered to each group. This analysis yielded correlations of .74 between the 5-minute halves of USES Test B (Group II) and .71 between the 5-minute halves of the CSC test (Group I), both higher than the correlations of .66 between scores on the first 5-minute segments of two 10-minute tests.

The original method of determining 5-minute reliability used in this study (separated 5-minute test segments) might seem the more rigorous, because two different test forms were involved. However, on the USES forms, most examinees had gone only about halfway through the text at the 5-minute point and were typing fresh material during the second 5 minutes. The correlation between these 5-minute segments is almost the same as alternate form reliability. On the CSC form, which was only about half as long as the USES forms, most examinees were repeating the material during the second 5 minutes, producing a test-retest situation. The fact that the CSC form reliability for the successive 5-minute segments was no higher than the same reliability for the USES form indicates that a test-retest reliability is not spuriously high.

## V. Conclusions

### A. Comparability

This study showed that both speed and error scores obtained on the USES typing test are comparable to those obtained on the CSC test, which has different content, format, length and time limit.

Some generalization of these results beyond the specific tests used in this study is proper: it is likely that other tests with one or more of the variations which existed in this study would also yield comparable scores. (Error scores on tests with different time limits must, of course, be adjusted to reflect the time limit if they are to be directly comparable.) However, if the variations were much greater than those in this study (for example, a 5-minute versus a 20-minute time limit) the comparability of resulting scores could not be taken for granted. It should also be borne in mind that the complexity of the text, as measured by the average number of strokes per word, was approximately the same for the two tests in this study. The effect on scores of differences in this variable was, therefore, not determined.

B. Reliability

For speed scores, the alternate form reliability of the 5-minute tests and the 10-minute tests were found to be substantially the same. For error scores, the reliability of the 10-minute tests exceeded that of the 5-minute tests by a significant margin when the reliability for the latter was computed between the first 5-minute segments of each 10-minute test. The difference between the 5- and 10-minute error score reliability coefficients was lessened considerably when the reliabilities for 5-minute tests were computed by correlating errors on the two successive 5-minute segments of the first test.